

A Model-Based Protocol for the Diagnosis of von Willebrand Disease

Christopher Castaldello^a, Federico Galvanin^b, Alessandra Casonato^c, Roberto Padrini^c,
Massimiliano Barolo^a and Fabrizio Bezzo^{a*}

^aCAPE-Lab – Computer-Aided Process Engineering Laboratory,
Department of Industrial Engineering, University of Padova, via Marzolo 9, 35131 Padova, Italy

^bDepartment of Chemical Engineering, University College London,
Torrington Place, London WC1E 7JE, United Kingdom

^cDepartment of Medicine, University of Padua Medical School,
Via Ospedale Civile 105, 35128 Padova, Italy

*Contact: E-mail fabrizio.bezzo@unipd.it, Tel. +39 049 827 5468

ABSTRACT

Von Willebrand disease (VWD) is one of the main inherited coagulation disorder. It is caused by a deficiency and/or a dysfunction of the von Willebrand factor (VWF), a fundamental multimeric glycoprotein involved in the hemostasis process. Correct detection of the disease is not an easy task because the disease manifests itself in many variants and a high intra-subject variability is observed. For these reasons, the diagnostic clinical trials typically rely on a 24-hour sampling protocol, which makes the overall test long, stressful and costly. Using a new pharmacokinetic model derived from Galvanin et al. (*AIChE Journal*, 2014, 60, 1718-1727) this study aims at: i) assessing the theoretical possibility to perform a shorter clinical test, ii) proposing a set of model-based diagnostic methods as a support for the clinical team.

A preliminary information analysis is performed in order to understand which sampling instants are more informative for model identification. This allowed us to propose a novel, 8-h diagnostic protocol which is still able to ensure model identifiability. Three alternative diagnostic methods are then proposed based on this short-length clinical protocol. One of them directly uses the pharmacokinetic model, whereas the other two are based on the use of three indices (two pharmacokinetic indices, namely clearance, total VWF released and as third index the basal multimer ratio) to formulate the diagnosis problem as a classification one. The classification problem is then solved using K -nearest neighbors and linear discriminant analysis. Results show the theoretical feasibility of a VWD diagnosis based on a shorter protocol.

Keywords: von Willebrand disease, model identification, disease diagnosis, discriminant analysis, nearest neighbors.

INTRODUCTION

Von Willebrand disease (VWD) is the most widespread inherited bleeding disorder that leads to coagulation problems. Epidemiological studies in Italy and in the USA estimate the affected population in about 1%.^[1;2] The disease is caused by a deficiency and/or a dysfunction of the von Willebrand factor (VWF), a multimeric glycoprotein present in the bloodstream. VWF acts as carrier and stabilizer for the coagulation factor FVIII and shows its relevance during the hemostasis process since it promotes the adhesion, activation and aggregation of platelets on the damaged endothelium site, leading to the formation of the blood cloth.^[3;4] The main clinical symptoms of VWD include nosebleeds, bleeding from small lesion in skin, mucosa or the gastrointestinal tract, menorrhagia and excessive bleeding after trauma, surgical interventions or childbirth.^[5] In healthy people, VWF level in blood strongly depends on the ABO group, in particular the O group has lower VWF level with respect to non-O group.^[6] VWD can be divided into three main types, depending on the defect on the VWF: partial quantitative deficiency (Type 1), qualitative dysfunction (Type 2), total qualitative deficiency (Type 3). A special case is represented by Vicenza subtype,^[7] that presents both quantitative and qualitative defect on the multimers. Furthermore, Type 2 presents variants such as 2A, 2B, 2M and 2N.^[8]

Classic diagnosis of the disease relies on the protocol that is based on the subcutaneous injection of DDAVP (1-desamino-8-D-arginine vasopressin), which forces VWF release, followed by the collection of 10 blood samples over a total time of 24 h (1440 min), which is in fact an invasive and stressful situation for the patients, as well as expensive in terms of nurse assistance. The test is complex and can be carried out only in specialized clinical centers. Correct diagnosis is not an easy task due the variety of subtypes, poor reproducibility of measurements, significant inter-individual variability, and strongly relies on the experience and competence of medical doctors. For such reasons, a diagnosis methodology based on a shorter clinical test and supported by tools for the identification of the disease would be highly beneficial in terms of cost reduction and alleviation of a subject's stress.

This conceptual design study is meant to address the above issues. The approach relies on a pharmacokinetic model derived from Galvanin et al.^[9] simulating VWF behaviors in the bloodstream after DDAVP administration, and aims at assessing:

- the possibility to reduce the total time of the current DDAVP test, by confining it to 8 h (i.e., one shift);
- the feasibility of adopting such shorter test to assist the diagnosis of the disease.

The methodology may represent a first step towards the definition of a model-based protocol supporting practitioners in the diagnosis of VWD and its classification.

The article is structured as follows. First, it will be demonstrated through an information analysis that a shorter clinical test can be envisaged. Then, we will exploit the pharmacokinetic model to propose a procedure for diagnosis. Three methods will be considered, one derived from the method proposed by Galvanin et al.,^[9] and two methods relying on some widely known multivariate classification methods, namely the K -Nearest Neighbors (KNN) and Linear Discriminant Analysis (LDA) methods. Finally, the results obtained from simulated case studies will be discussed.

THE PROPOSED DDAVP CLINICAL TEST

In this section, after illustrating the available clinical data obtained from real patients and the key features of the pharmacokinetic model, the possibility to propose a shorter clinical test is assessed.

Available Data

Experiments were performed according to the Declaration of Helsinki for respecting the human research ethics. Data are available for four classes of subjects: HO, HnonO, 2B and Vicenza. While the first two classes refer to healthy subjects of different ABO group, the latter two include subjects affected by different VWD subtypes. The protocol used to obtain the data consists in DDAVP subcutaneous administration ($0.3 \mu\text{g}/\text{kg}$ of body weight), followed by 10 blood sampling at fixed times (0, 15, 30, 60, 120, 180, 240, 360, 480, 1440 min), for a total time of 24 h.^[10] Two assays are performed on each sample:^[8]

- von Willebrand factor antigen (VWF:Ag), which is an immunoassay that measures the total concentration of VWF protein in the plasma.
- von Willebrand factor collagen binding (VWF:CB), which is an assay that measures binding of VWF multimers on collagen of the vascular site. Since the assay is size-dependent, only high molecular-weight multimers are measured by this test due to their higher affinity.

The results of both assays are expressed as International Unit per deciliter (IU/dL).

In addition to VWF:Ag and VWF:CB data at fixed times, the body weight BW of each patient is necessary. Most clinical data are the same as the ones used and validated in Galvanin et al.;^[9] additional data were incorporated after removing outliers that could not be explained in terms of measurement errors.

In order to obtain a qualitative and quantitative profile for each measured response, a fictitious “average” subject for each class was created: each historical average subject represents the average value for each class at every sampling time among all available subjects. The adoption of an historical subject allows one to obtain the class-dependent data profile of Figure 1, where uncertainty is expressed as standard error (SE).

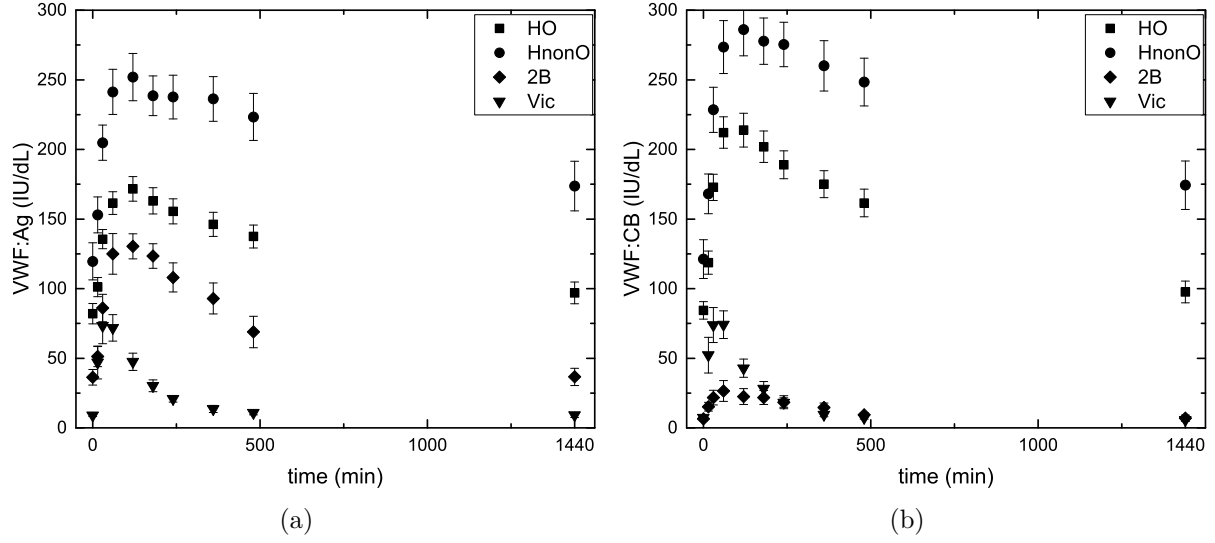


Figure 1: Historical average subject data for each class for responses of (a) VWF:Ag and (b) VWF:CB. Error bars represent SE of the mean.

The Pharmacokinetic Model

The availability of a reliable pharmacokinetic model is important for the correct evaluation of a new experiment throughout in-silico analysis. After DDAVP administration, three sequential events occur: release of VWF from endothelial cells, proteolysis of higher multimers into lower multimers (thanks to the ADAMTS13 enzyme) and clearance from blood streams.^[11;12]

Galvanin et al. proposed a three-compartments model describing the three events above.^[9] However, its complexity makes model identification difficult if based on a limited number of samples, as occurs in conventional clinical tests.^[13]

For this reason, a new simplified model is used in this study, which describes only the main chemical pathways involved. This model (Figure 2) is composed by two compartments and relies on the following assumptions:

- at the basal state, only H (high molecular-weight) and L (low molecular-weight) multimers are present;
- release of UL (ultra large molecular-weight) and of additional H multimers is a consequence of DDAVP administration;
- elimination constant k_e is considered independent from multimer size and it is the same for both compartments;^[14]
- VWF:Ag test measures the overall amount of UL, H and L multimers;
- VWF:CB test measures the amount of UL+H multimers.

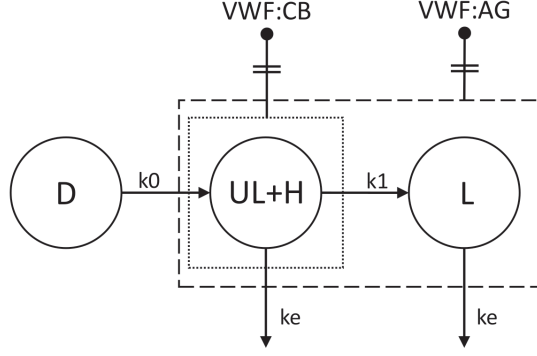


Figure 2: Structure of the new pharmacokinetic model, showing the compartments involved and the available measurements.

The model is described by the following set of differential and algebraic equations:

$$\frac{dx^{\text{UL+H}}}{dt} = k_0 D e^{-k_0(t-t_{\max})} - k_1(x^{\text{UL+H}} - x_b^{\text{UL+H}}) - k_e(x^{\text{UL+H}} - x_b^{\text{UL+H}}) \quad (1)$$

$$\frac{dx^{\text{L}}}{dt} = k_1(x^{\text{UL+H}} - x_b^{\text{UL+H}}) - k_e(x^{\text{L}} - x_b^{\text{L}}) \quad (2)$$

$$x_b^{\text{UL+H}} = y_b^{\text{CB}} V_d \quad (3)$$

$$x_b^{\text{L}} = y_b^{\text{Ag}} V_d - x_b^{\text{UL+H}} \quad (4)$$

$$y^{\text{Ag}} = \frac{x^{\text{UL+H}} + x^{\text{L}}}{V_d} \quad (5)$$

$$y^{\text{CB}} = \frac{x^{\text{UL+H}}}{V_d} \quad (6)$$

$$y^{\text{CB1}} = k y^{\text{CB}} \frac{y_b^{\text{Ag}}}{y_b^{\text{CB}}} = k y^{\text{CB}} \left(1 + \frac{x_b^{\text{L}}}{x_b^{\text{UL+H}}} \right) \quad (7)$$

where $x^{\text{UL+H}}$ and x^{L} represent respectively the number of UL+H and L multimer units of VWF present in the bloodstream, the subscript b indicates each variable at their

basal condition, whereas k_0 and D are release parameters, t_{\max} is the time of maximum response (also called lag time), and k_1 and k_e are parameters related respectively to proteolysis and clearance phases. Equation (7) is adopted to take into account the different affinity observed by different VWD subtypes with respect to collagen binding test. In the calculations of the basal multimers number (Equations (3) and (4)) and in the conversion from multimer unit into concentrations (Equations (5) and (6)), the volume of distribution is approximated using the body weight BW , following Menache et al.:^[15]

$$V_d = 0.4 BW. \quad (8)$$

A reparametrisation was exploited in order to achieve model identifiability.^[16] In this study, the commercial software gPROMS[®] (by Process Systems Enterprise, Ltd.) was used for parameter estimation. The final parameters vector $\boldsymbol{\theta}$ is:

$$\boldsymbol{\theta} = [k_0 \ k_1 \ k_e \ D \ k \ y_b^{\text{CB}} \ D/t_{\max}] . \quad (9)$$

In addition to the above equations, three indices can be calculated that can be used to improve the understanding of the key events involving VWF in bloodstream:

$$CL = k_e V_d \quad (10)$$

$$Q = \frac{1}{BW} \int_0^\tau k_0 D e^{-k_0(t-t_{\max})} dt \quad (11)$$

$$R = \frac{\text{VWF:CB}_b}{\text{VWF:Ag}_b} \quad (12)$$

where τ is the test duration. CL [mL/kg/h] (clearance) is related to elimination of VWF while Q [IU/kg] is total amount of VWF released; both of them are pharmacokinetic indices derived from the model identification, whereas R [–] (the ratio of VWF:CB and VWF:Ag at the basal state) is calculated directly from experimental data. These class-dependent indices will be exploited with classification methods in order to assist the diagnosis.

Parameter estimation

Parameter estimation was performed on each historical average subject data in a similar way as in Galvanin et al.,^[17] in such a way as to predict class-dependent profiles. The lack of fit test (LOF) is used to assess the fitting of the estimated responses and it is

based on the sum of the squared weighted residuals:

$$\chi^2 = \sum_{i=1}^{N_y} \left[(\mathbf{y}_i - \hat{\mathbf{y}}_i) \boldsymbol{\Sigma}_{y_i}^{-1} (\mathbf{y}_i - \hat{\mathbf{y}}_i) \right] \quad (13)$$

where \mathbf{y}_i is the experimental data, $\hat{\mathbf{y}}_i$ is the predicted value, $\boldsymbol{\Sigma}_{y_i}$ is the matrix of the measurements errors, with all of three referred to the i -th response. This calculated value is compared to the one taken as reference and retrieved from a χ^2 distribution with $N_{\text{sp}} - N_{\theta}$ (number of sampling points and number of parameters to be estimated, respectively) degrees of freedom and with a confidence level of 95%. If the χ^2 value obtained at the end of the parameter estimation is lower than the reference one, then the estimated model is a good representation of the clinical data.

Parameters estimation results for these subjects are reported in Table 1 and show that, for all subjects a statistically satisfactory identification of the parameters is achieved even if good fitting, through lack of fit test, is not always obtained. This is due to high variability within the class, especially for healthy subjects.

These parameters values provide a representative predicted profile for each class. They will be used later as initial guesses when a new identification exercise is carried out as a step of the diagnosis procedure.

Table 1: Results of parameter estimation task for each historical average subject of the class. The cases in which lack of fit test failed are highlighted in bold.

parameter	HO	HnonO	2B	Vic
k_0	0.0265	0.0287	0.0146	0.0403
k_1	0.0004	0.0001	0.0034	0.0012
k_e	0.0013	0.0007	0.0036	0.0099
D	141.2627	169.9747	183.7347	231.1436
k	0.9864	0.9688	0.1849	0.7307
y_b^{CB}	49.9723	84.6426	24.5367	5.8837
D/t_{max}	1.2799	1.5594	0.8550	3.5216
$\chi^2(\chi_{\text{Ref}}^2)$	168.2 (26.3)	192.3 (26.3)	97.8 (26.3)	15.9 (26.3)

Developing a Shorter Protocol

We now describe the procedure used to propose a new shorter clinical test and the assessment of its feasibility, with the aim of maintaining model identifiability.

In-silico tests

The procedure used to obtain new in-silico data is outlined in Figure 3, and it is composed by three main steps:

- at first, individual parameter identification is performed, using clinical data (VWF:Ag,

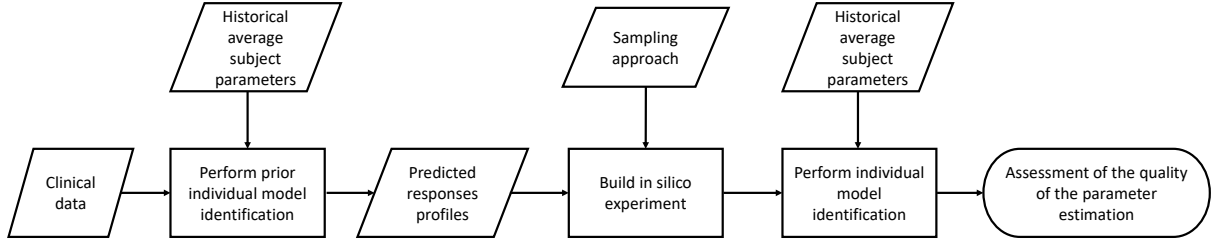


Figure 3: Three-step procedure for in-silico experiment assessment.

Table 2: Values of σ_{err} under the assumption of constant variance model for the generation of errors, for each class of subjects. These values were calculated and rounded to the closest integer, from the standard deviation of the data of the historical average subject with respect to the predicted responses profiles.

	HO	HnonO	2B	Vic
σ_{err} (IU/dL)	4	5	2	2

VWF:CB and BW) together with parameters available from the historical average subject of the membership class as initial guesses;

- using the predicted profiles of the responses (obtained from the previous step) and the assumed sampling schedule, an in-silico experiment is built adding the error representing the variability of the measures of the subject. The error used to build in-silico data is generated randomly following a Gaussian distribution with the class-dependent standard deviation σ_{err} reported in Table 2;
- the new in-silico experiment is used to carry out an additional parameter identification exercise.

At the end of the procedure, statistics such χ^2 are calculated, assessing the quality of the parameter estimation for each subject. It should be pointed out that the use of a correct error model impacts on the results of the estimation. In fact, the σ_{err} reproduces within a subject the average intra-variability of each class.

Selecting the sampling approach

The discrete form of the dynamic Fisher Information Matrix (FIM) \mathbf{H}_θ , for a series of N_{sp} samples, is exploited to assess the information carried by an experiment:^[18]

$$\mathbf{H}_\theta^*(\boldsymbol{\theta}, \boldsymbol{\varphi}) = \sum_{i=1}^{N_y} \sum_{j=1}^{N_y} s_{ij} \mathbf{Q}_i^T \mathbf{Q}_j \quad (14)$$

$$\mathbf{H}_\theta(\boldsymbol{\theta}, \boldsymbol{\varphi}) = \sum_{k=1}^{N_{\text{sp}}} \mathbf{H}_\theta^*(\boldsymbol{\theta}, \boldsymbol{\varphi}_k) \quad (15)$$

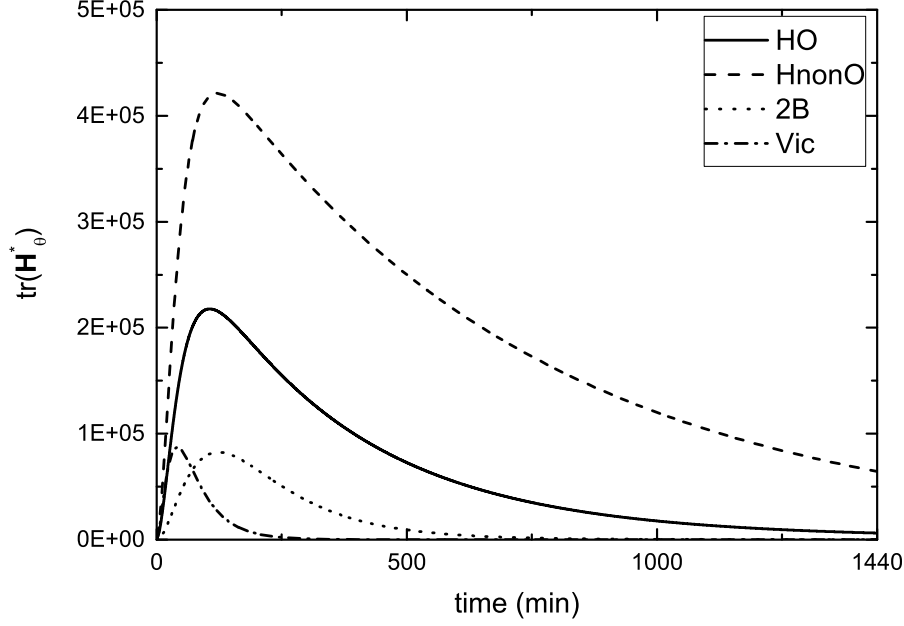


Figure 4: Profiles of the trace of the information matrix over time, computed for each historical average subject.

where

- \mathbf{Q}_i is the $[N_{\text{sp}} \times N_{\theta}]$ matrix of dynamic sensitivity for the i -th measured response;
- s_{ij} is the i, j -th element of the inverse matrix of the Σ_y variance-covariance matrix of the experimental measures;
- while $\boldsymbol{\varphi}$ is the global design vector, $\boldsymbol{\varphi}_k$ is the design vector involved in the construction of each sampling point.

The criterion adopted for the evaluation of the amount of information is the so called A-optimal,^[19] and was chosen because of its ease of implementation, based on the calculation of the trace of the dynamic information matrix. The maximum value of the trace of the information matrix corresponds to the maximum obtainable information. The profiles of dynamic trace $\text{tr}(\mathbf{H}_{\theta}^*)$ for each historical average subject are shown in Figure 4: it is therefore possible to identify the time interval within which samples provide the larger amount of information. In particular, it is clear that the most informational time interval is represented by the first 500 minutes of the test, within which the information peak for each subject class is always located.

Equation (15) allows calculating the total information $\text{tr}(\mathbf{H}_{\theta})$ for a discrete sampling, as would occur in a clinical test. Two sampling approaches based on 10 and 15 evenly spaced samples are considered (15 is the maximum number of samples that seem viable considering the time required for the analysis and typical nurse staff availability). Different durations of the clinical test are assessed. As an example, Figure 5 shows the total information that can be collected from the historical average subject of the HnonO class; results from

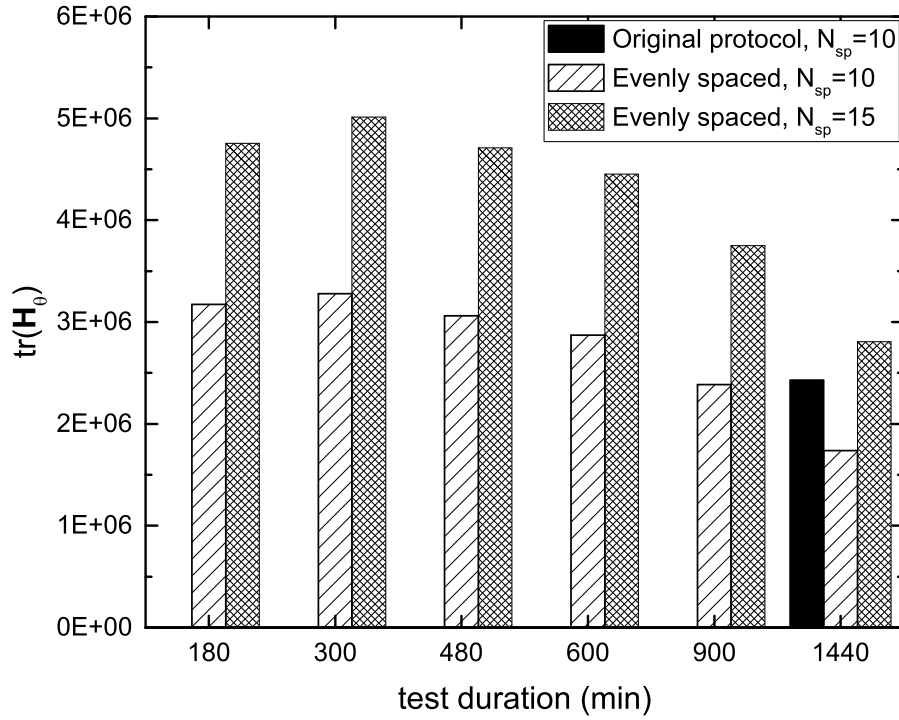


Figure 5: Total amount of information for different sampling approaches, different number of samples and different test duration. The results refer to the HnonO historical average subject.

other classes are qualitatively similar. It can be observed that the information peak is reached for a test duration of 300 min (i.e., almost 5 times shorter than current practice); a further time reduction would cut off a significant informative interval. Obviously the more the sampling points, the more information can be collected.

From this preliminary analysis, it appears that the reduction of the duration of the clinical test is theoretically possible, since the information amount can be improved by sampling the system in more favorable time instants. A clinical test based on 15 evenly spaced samples over 480 min (i.e., 3 times shorter than current practice) was eventually considered. Although the amount of information that would be obtained from a 300 min test is theoretically greater, the sampling frequency would be excessively high in practical terms (on the other hand, reducing to 10 the number of samples would determine a significant loss of information). It must be noted that 15 samples in 480 min represents a slightly worse condition for the patient with respect to the current protocol; nevertheless, the protocol remains feasible and represents a first attempt to identify a procedure that could be carried out in a medical facility.

Time reduction results

The newly devised short protocol was applied in-silico to the historical average subjects, and the model was found identifiable for all of them.

This allowed proposing some diagnostic methods, based on this reduced protocol. Once

model identification for each individual subject to be diagnosed is performed, then the pharmacokinetic indices CL and Q are calculated for each subject following Equations (10) and (11). A multiple shooting technique is adopted to reduce the effect of local optima in the parameter estimation algorithm: initial guesses for the parameters are taken from each average subject in Table 1. Eventually, the best fitting (evaluated through LOF) accross all shootings is considered. These two forenamed indices will be used in the diagnostic techniques discussed in the next sections.

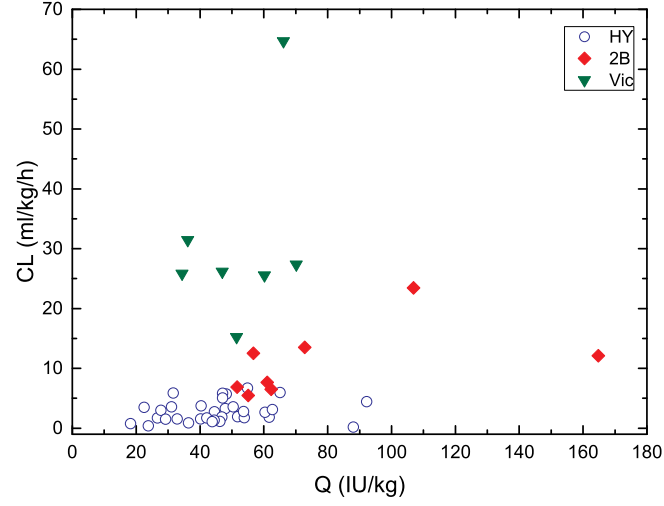
The three retrieved indices CL , Q and R are represented graphically using three 2D plots in Figure 6, to best highlight the behavior of each class. These plots meet medical findings,^[8] e.g. R index for 2B class is known to be usually smaller than in healthy subjects due to an increased proteolytic degradation of high molecular-weight multimers. Also, CL is greater for classes 2B and Vic with respect to healthy subjects; namely Vic subjects present a significantly higher clearance, whereas for 2B subjects the effect is still present but less marked. From a visual point of view it appears that class separation can be achieved quite neatly using two indices only (CL vs R ; Figure 6(b)). However, since the entire CL , Q and R triplet represents the individual behavior of a given subjects, all three indices will be used, as discussed later with reference to the KNN method.

A PROCEDURE FOR DIAGNOSING VWD

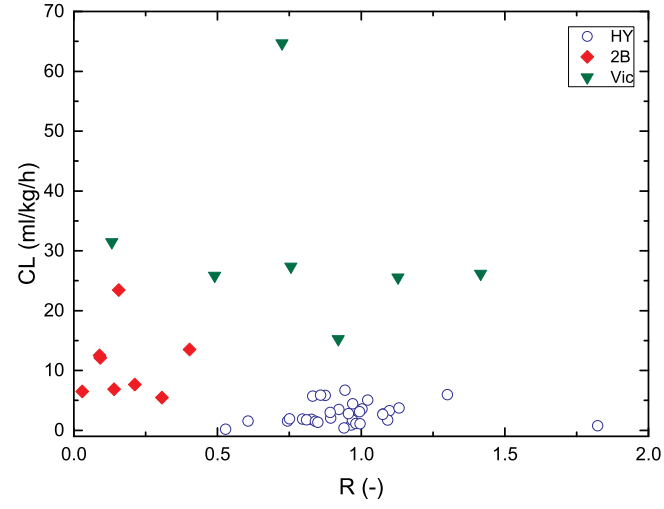
Three diagnosis methodologies are now proposed and tested. In particular Method 1 is derived from the methodology defined by Galvanin et al.,^[9] whereas Methods 2 and 3 are based on well-known classification methods, namely KNN and LDA. The main objective is to distinguish healthy subjects from VWD-affected ones, and in the latter case assess the VWD subtype.

Method 1 is the more computationally demanding and is based on a classification among HO, HnonO, 2B and Vic classes, for a better characterization of the uncertainty of each class. The method capability of correctly classifying healthy O and nonO subjects is not assessed here: as previously pointed out, the target is to distinguish between healthy and VWD-affected subjects. Methods 2 and 3 are based on two classifications among HY, 2B and Vic classes (HY is obtained by merging HO and HnonO classes). In particular, Method 2 is based on three indices (R , Q and CL). Method 3 relies on CL and R only; hence, it represents the simplest strategy for diagnosis and allows for an intuitive graphical classification. The general idea is that a combined use of all three methods can assist the medical staff in the diagnosis (i.e., classification) job.

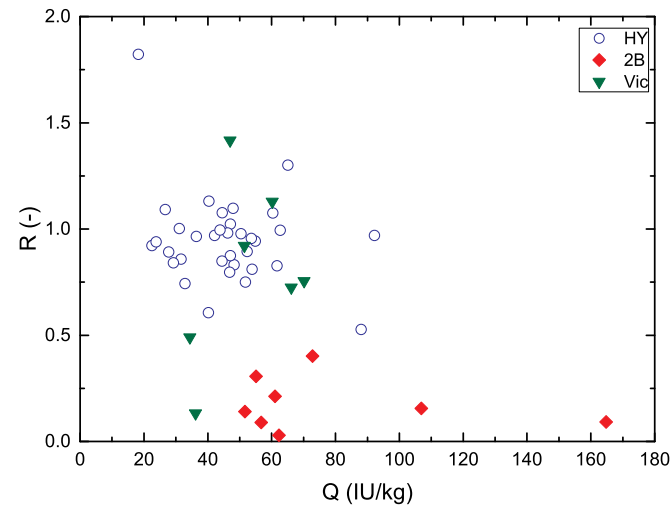
In the following, in-silico clinical tests are used to assess the diagnosis potential of the three methods.



(a)



(b)



(c)

Figure 6: Pairwise representation of the three indices Q , CL and R , characterizing a given subject, for all subjects considered in this study. Subjects are grouped into three classes thanks to medical examination: healthy (HY), 2B and Vic.

Method 1: SSWR Diagnosis

The method proposed by Galvanin et al. is applied.^[9] This procedure is based on an individualized identification approach that couples two source of the VWF:Ag and VWF:CB data: the individual subject, and the historical average subject of a class together with its class variance. The identification task is therefore executed as many times as the number of the classes, adopting in each case the corresponding historical average subject and its class variance. The key difference with respect to the original study of Galvanin et al. is that here a simplified pharmacokinetic model is used and that a shorter protocol for the clinical test is applied. A heteroscedastic variance model calculated from experimental data is adopted to calculate class variance:

$$\sigma_e^2 = \omega^2(y^2)^\gamma \quad (16)$$

where the two variance parameters ω and γ are obtained from a least squares regression from the average data (Table 3). Thanks to these class-dependent variance parameters, the description of the class-dependent population uncertainty for VWF:Ag and VWF:CB measured responses over time is possible. It must be noted that the so-built variance changes significantly between classes, since it depends also from the measured responses, as already seen in Figure 1.

Table 3: Heteroscedastic parameters for the four classes considered in classification Method 1. The parameters were obtained from regression on real data.

Response	Param.	HO	HnonO	2B	Vic
VWF:AG	ω	7.917	20.589	4.418	1.252
	γ	0.325	0.210	0.397	0.762
VWF:CB	ω	2.436	17.090	0.128	2.028
	γ	0.578	0.258	1.533	0.653

Following the approach described in Galvanin et al.,^[9] the diagnosis is obtained via SSWR index (sum of squared weighted residuals) calculated after each parameter estimation task. SSWR represents the closeness of the subject to a specific class, using the following equations:

$$\text{SSWR} = \text{SSWR}_{\text{subj}} + \text{SSWR}_{\text{HAS}} \quad (17)$$

$$\text{SSWR}_{\text{subj}} = \sum_{y=1}^{N_y} \sum_{\text{sp}}^{N_{\text{sp}}} \frac{(y - \hat{y})^2}{\sigma^2} \quad (18)$$

$$\text{SSWR}_{\text{HAS}} = \sum_{y} \sum_{N_{\text{sp}}} \frac{(\bar{y} - \hat{y})^2}{\sigma_e^2} \quad (19)$$

where y is the experimental data of the unknown subject, \hat{y} is the predicted value of the response, \bar{y} is the predicted response of historical average subject for each class. It must be pointed out that:

- $\text{SSWR}_{\text{subj}}$ represents the similarity between the experimental points and the predicted profiles of the subject, weighted with the standard deviation of the measure (fixed at $\sigma = 2 \text{ IU/dL}$);
- SSWR_{HAS} represents the similarity between each historical average subject and the predicted profiles, weighted with the heteroscedastic variance model σ_e .

The subject is assigned to the class for which the smaller value of SSWR is found (i.e., the class that is more similar to the response of the subject).

Method 2: K -Nearest Neighbors Diagnosis

Classification through KNN is one of the most used and simple classification techniques in multi-variables problems, and has been successfully used in a variety of medical applications, including cancer detection.^[20]

KNN allows classifying an unknown subject according to the K known subjects who are closest to him/her. From a computational point of view, the distance from each training (known) subject is calculated, then subjects are ordered in ascending order of distance and the K closest subjects are selected.^[21] Distances of all training subjects are then weighted to give more importance to closest neighbors.^[22] Eventually, weights are normalized by class so their sum is equal to the prior probability (assumed uniform in this study), for each class.

In this study the distance between two subjects is calculated following the cosine distance as defined in Equation (20), and it is weighted by a squared inverse law:

$$d_{\cos} = 1 - \frac{\mathbf{u}\mathbf{v}}{\sqrt{(\mathbf{u}\mathbf{v}')(\mathbf{v}\mathbf{v}')}}. \quad (20)$$

where \mathbf{u} is the unknown subject vector and \mathbf{v} is a training vector, both containing the indices CL , Q and R used as predictors.

A critical step in this method is assigning the optimal value of K , especially if classes are not of the same size.^[23] In this study, $K = 3$ was selected by a trial and error, and taking into account the small number of available VWD subjects. The healthy O and healthy non-O subjects were merged into a single class (denoted by HY), and therefore only three classes were considered when applying this method (namely HY, 2B and Vic).

The classification task acts in minimizing the expected classification cost, obtaining the predicted class \hat{l} :

$$\hat{l} = \arg \min_{l=1, \dots, N_G} \sum_{N_g=1}^G \hat{P}(g|\mathbf{u}) C(l|g) \quad (21)$$

where N_G is the number of classes and $\hat{P}(g|\mathbf{u})$ is the posterior probability for subject \mathbf{u} of belonging to class g . The posterior probability is calculated among the subjects included in the K nearest neighbors, as:

$$\hat{P}(g|\mathbf{u}) = \frac{\sum_{i \in K} W(i) \mathcal{Y}(i)}{\sum_{i \in K} W(i)} \quad (22)$$

where $W(i)$ is the normalized weight of the subject i and $\mathcal{Y}(i)$ is a factor that is equal to one if subject i belongs to class g , and zero otherwise. In order to achieve the best separation between healthy and non-healthy classes, a misclassification cost $C(l|g)$ was used in Equation (21), representing the ability to avoid subject misclassification to class l while its true class is g .^[24] For every couple of classes, the misclassification cost is assigned a-priori and $C(l|g)$ could be different with respect to $C(g|l)$. Trivially, the misclassification cost for a class with respect to the same class is zero, since there is no misclassification. The default value for the misclassification cost is 1, meaning that no cost is assigned: this case is taken as base case to assess the method performance. The misclassification costs $C(\text{HY}|\text{2B})$ and $C(\text{HY}|\text{Vic})$ are here tuned up with a trial and error procedure to avoid false negative (e.g., a 2B or Vic subject assigned erroneously to HY class), thus making the model more sensitive to VWD detection.

Applying KNN classification, a leave-one-out procedure was used to check the predictive capability of the method. Practically, an item is cyclically removed from the calibration of the model and then the same item is predicted. At each loop the model is calibrated again. At the end of the loops, when all subjects are classified, performances indices can be calculated.

Assessing performance of the classification

The classification performances are assessed with some a posteriori indices. At the basis of the performance evaluation there is the *confusion matrix*, a square matrix $[N_G \times N_G]$, in which each n_{gk} element is the number of subjects belonging to class g assigned in class k (note that diagonal elements n_{gg} represent correctly assigned subjects). The following performances indices are used:^[23]

- Error rate ER represents the percentage of wrongly assigned subjects:

$$ER = 1 - \frac{\sum_{g=1}^G n_{gg}}{n}. \quad (23)$$

where n is the total number of subjects.

- Sensitivity Sn_g is the ability to recognize correctly subjects belonging to g -th class:

$$Sn_g = \frac{n_{gg}}{n}. \quad (24)$$

- Specificity Sp_g represents the ability of g -th class to reject subjects from other classes:

$$Sp_g = \frac{\sum_{g=1}^G (n'_k - n_{gk})}{n - n_g} \quad \text{for } k \neq g \quad (25)$$

$$n'_k = \sum_{g=1}^G n_{gk} \quad (26)$$

where n'_k is the total number of subjects assigned to k -th class.

- Precision Pr_g is the ability not to include subjects from other classes, as the ratio between correctly assigned subjects and total subjects assigned to that class:

$$Pr_g = \frac{n_{gg}}{n'_g}. \quad (27)$$

For ER the best value is 0, whereas for the others the best value is 1.

Method 3: Linear Discriminant Analysis Diagnosis

Discriminant analysis is also a widely adopted classification method. It has been used, e.g., as diagnostic method in Parkinson detection,^[25] or for genetic classification using FVIII clinical tests in VWD.^[26] Its objective is to find a direction in the data space that allows the maximization of the between-class variance and the minimization of the within-class variance;^[27] these directions are called discriminant functions.

In particular, Linear Discriminant Analysis (LDA) relies on the hypothesis that all the g classes generate data with a Gaussian density function $f_g(\mathbf{x})$ (where x is the generic variable), with the same covariance matrix Σ but different means μ_g . With this assumption, $\hat{P}(g|\mathbf{u})$ is the posterior probability that a subject \mathbf{u} belongs to class g , which is calculated

knowing the prior probability π_g (assumed as uniform) and a normalization factor:

$$f_g(\mathbf{u}) = \frac{1}{(2\pi)^{1/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{u} - \boldsymbol{\mu}_g)^T \boldsymbol{\Sigma}^{-1}(\mathbf{u} - \boldsymbol{\mu}_g)\right] \quad (28)$$

$$\hat{P}(g|\mathbf{u}) = \frac{f_g(\mathbf{u})\pi_g}{\sum_{l=1}^G f_l(\mathbf{u})\pi_l} \quad (29)$$

The boundary between two classes is defined setting the log-ratio among the two posterior probabilities equal to zero (this defines the locus where posterior probability for two classes is the same): this leads to a linear function in \mathbf{x} (Equation (30)), whose coefficient can be easily retrieved, in this case a_0 is the prior probability and b_0 is the constant part of the Gaussian distribution ratio and the priors ratio.^[28]

$$\begin{aligned} \log\left(\frac{\hat{P}(g|\mathbf{x})}{\hat{P}(l|\mathbf{x})}\right) &= \log\frac{\pi_g}{\pi_l} - \frac{1}{2}(\boldsymbol{\mu}_g + \boldsymbol{\mu}_l)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_g - \boldsymbol{\mu}_l) + \mathbf{x}^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_g - \boldsymbol{\mu}_l) = 0 \\ &= a_0 + b_0 + b_1 x_1 + b_2 x_2 = 0 \end{aligned} \quad (30)$$

The class prediction is based on the minimization of the expected classification cost in the same way as in Equation (21), but with the posterior probability assumptions as in Equation (29). In LDA, the effect of the misclassification cost $C(l|g)$ on the boundary between two classes is clear, since it affects only the constant term of the linear function in Equation (30) shifting the involved boundary towards unwanted class, e.g. a $C(l|g) > 1$ shifts the boundary towards class l .^[24] In fact, among two generic classes I and II, the decision rule assigns subject \mathbf{u} to class I if $\hat{P}(II|\mathbf{u})C(I|II) < \hat{P}(I|\mathbf{u})C(II|I)$. Thus, from a graphical point of view, the subject \mathbf{u} belongs to class I if the following is true:

$$b_0 + b_1 u_1 + b_2 u_2 > \log \frac{\pi_{II}C(I|II)}{\pi_I C(II|I)} \quad (31)$$

in which the constant part is derived from the Gaussian density distribution. Subject is assigned to class II in the opposite situation.

Analogously to KNN, LDA was carried out dividing the available subjects into three classes (HY, 2B and Vic), with the intent of the recognition of unhealthy subjects. Following the same steps as KNN, an optimal misclassification cost was used in order not to misclassify subjects affected by VWD. This value is found again with a trial and error procedure, to meet the right trade-off between error rate and unhealthy subjects detection, increasing it from the default unitary value.

For the specific case of this work, only CL and R are used as predictors to simplify the graphical representation, thanks to the narrower overlapping area between classes observed in Figure 6(b), with respect to the other two plots.

Classification performances of LDA are assessed following a leave-one-out procedure, and by adopting the same performances indices as in Method 2.

RESULTS

Method 1 Results: SSWR Diagnosis

Similarly to the original reference,^[9] the method was tested on a limited number (four) of (supposed to be) unknown subjects. In order not to bias the method, the four randomly selected subjects were not included in the calculation of the historical average subjects and in the calculation of the heteroscedastic variance parameters.

Results are reported in Table 4, showing that Method 1 allows a clear identification between healthy and ill subjects. Also the diagnosis of the two subjects affected by VWD is achieved correctly. Some issues appear in the distinction between O and non-O healthy subjects, but as discussed earlier this is a minor issue in this work. Furthermore the SSWR values obtained for healthy subject coupled with VWD classes is much more higher than the values obtained from the unhealthy subjects when coupled with healthy classes, this is another factor that will help practitioners in the correct class assignment.

In conclusion, Method 1 is capable of producing a good classification performance, even when it is based on a simpler pharmacokinetic model. Furthermore, the approach appears to be compatible with a shorter (8 h) clinical test.

Table 4: Results of model-based diagnosis in terms of SSWR on four unknown subjects. Between parenthesis the real membership class, in boldface the lowest value suggesting the diagnosis.

Unknown subject (real class)	HO	HnonO	2B	Vic
I (HO)	116.3	185.4	1434.6	1366.0
II (HnonO)	187.5	241.8	5012.1	4425.4
III (2B)	183.0	235.9	162.1	682.1
IV (Vic)	295.5	315.8	221.1	36.1

Method 2 Results: K -Nearest-Neighbors Diagnosis

With regards to the misclassification cost, after a trial and error procedure a misclassification cost $C(\text{HY}|2\text{B}) = C(\text{HY}|\text{Vic}) = 5$ was selected.

The results of this classification analysis are displayed in Table 5, where a reference case with $C(\text{HY}|2\text{B}) = C(\text{HY}|\text{Vic}) = 1$ is used to clarify the influence of the cost parameter. In general, the classification correctly assesses over the 95% of the subjects (all the unknown subjects used in the Method 1 are correctly classified with KNN). Note that in the base case a Vic subject is classified as healthy (resulting in a $Sn_{\text{Vic}} < 1$) and a healthy subject is classified as 2B (observable from the $Sp_{2\text{B}}$ and $Sp_{2\text{B}}$ both less than 1).

After increasing the misclassification cost up to 5, ER decreases and no false negative classification is observed anymore, thus improving the predictive capability of Method 2 to detect unhealthy subjects and leading to a satisfactory classification performance where all unhealthy subjects are correctly identified (Sn_{2B} and Sn_{Vic} are both equal to one). One healthy subject is still misclassified as 2B and in fact, Sp_{2B} and Pr_{2B} indices are equal to the reference case.

Table 5: KNN classification. Results in terms of performance indices in the cases of no misclassification cost ($C(HY|2B) = C(HY|Vic) = 1$) and with a value of 5.

$C(HY 2B) = C(HY Vic)$	ER	Sensitivity (Sn)			Specificity (Sp)			Precision (Pr)		
		HY	2B	Vic	HY	2B	Vic	HY	2B	Vic
1	0.04	0.97	1.00	0.86	0.93	0.98	1.00	0.97	0.89	1.00
5	0.02	0.97	1.00	1.00	1.00	0.98	1.00	1.00	0.89	1.00

Method 3 Results: Linear Discriminant Analysis Diagnosis

After a trial and error procedure, a misclassification cost equal to $C(HY|2B) = C(H|Vic) = 10$ was selected, from the default value of 1. The results for LDA in terms of separation zones in the $CL-R$ plot are displayed in Figure 7, where also the effect of the misclassification cost is evident. Since misclassification costs as defined above act only between HY/2B and HY/Vic classes, only those boundaries are shifted from the base case. Consistently to the intent of making the Method more sensitive to VWD detection, the effect of higher misclassification costs is that of increasing the size of the region comprising unhealthy subjects.

Observing the classification results in Table 6, the correct diagnosis is achieved in approximately the 95% of the cases and the four unknown subjects used in the Method 1 are correctly assigned. Similarly to Method 2, if no misclassification cost is assigned (reference case), i.e. $C(HY|2B) = C(HY|Vic) = 1$, one subject (the same subject as in Method 2) is classified as HY (instead of Vic), and another subject is classified as 2B (instead of Vic). By increasing the misclassification costs up to 10, no false negative is observed anymore. However, as a side-effect, two HY subjects are assigned to 2B class (lower Sn_H with respect to the base case), and one Vic subject is still classified as 2B. This impacts on the Sp_{2B} and Pr_{2B} values, which are lower than in the reference case, and on the ER value, which is larger.

In this classification method, it was also possible to investigate the effect of the model calibration set. Following a leave-one-out procedure, it was possible to identify a region of uncertainty, within which the classification bounds vary according to the calibration set (Figure 8). Subjects falling within the uncertainty region may be misclassified, since their classification depends on the specific set of data.

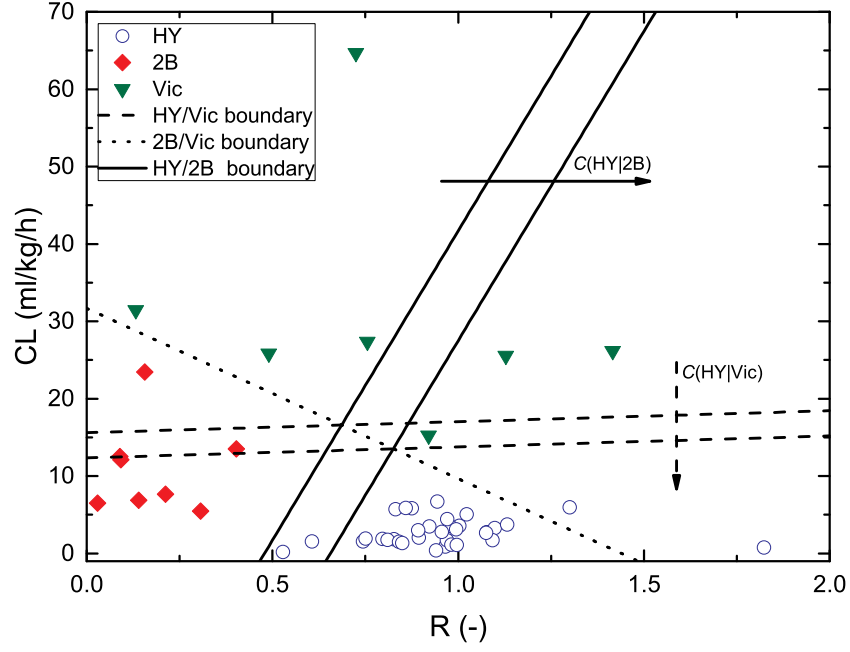


Figure 7: LDA classification. Representation of boundaries between each couple of classes obtained with all subjects used in model calibration. In this figure is clear the effect of misclassification costs $C(HY|2B)$ and $C(HY|Vic)$: increasing them (the direction of the arrows), the bounds are shifted towards healthy region. Since $C(2B|Vic)$ was not modified, this boundary is not affected.

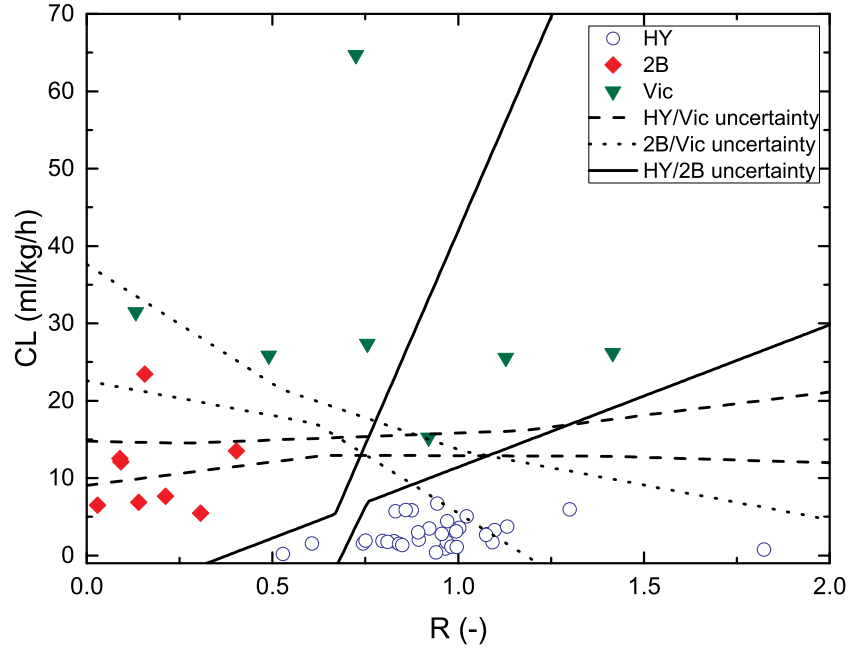


Figure 8: LDA classification. Representation of uncertainty boundaries built with leave-one-out procedure with a $C(HY|2B) = C(HY|Vic) = 10$.

Table 6: LDA classification. Results in terms of performance indices in the cases of no misclassification cost ($C(\text{HY}|\text{2B}) = C(\text{HY}|\text{Vic}) = 1$) and with a value of 10.

$C(\text{HY} \text{2B}) = C(\text{HY} \text{Vic})$	ER	Sensitivity (Sn)			Specificity (Sp)			Precision (Pr)		
		HY	2B	Vic	HY	2B	Vic	HY	2B	Vic
1	0.04	1.00	1.00	0.71	0.93	0.98	1.00	0.97	0.89	1.00
10	0.06	0.94	1.00	0.86	1.00	0.93	1.00	1.00	0.73	1.00

CONCLUSIONS

The results of this analysis have shown that by using a suitable pharmacokinetic model it is possible to envisage a model-based procedure for VWD diagnosis relying on a shorter clinical test than the one currently adopted in the medical practice (8 hours vs. 24 hours). The approach is based on three different Methods, built on different classification techniques and exploits the individuals VWF levels and the body weights of subjects. Simulated case studies have demonstrated the potential of the methodology, which appears to be quite robust in terms of detection of VWD-affected subjects, and which in most cases allows for the correct classification of the VWD typology.

Having said that, it should be recognized that the proposed methodology still has some limitations and cannot substitute the expertise of clinical practitioners. For instance, it is forbidden to administer the DDAVP test in cases of type 2B VWD in patients lacking in large VWF multimers:^[29] clearly, only a clinician can identify such situations and decide about the most convenient tests.

Future work, will aim at assessing the proposed approach via real clinical tests. If clinical data confirm the satisfactory results of this concept design, the successive step will be that of implementing a user-friendly tool to help clinicians in the characterization of the VWD.

ACKNOWLEDGEMENTS

Financial support from University of Padova through project 2015 “Development of physiological models to support the diagnosis of von Willebrand disease” is gratefully acknowledged.

NOMENCLATURE

Acronyms

ABO blood group classification system.

ADAMTS13 metalloprotease enzyme (A Disintegrin-like And Metalloprotease domain with ThromboSpondin type I motifs, member 13).

DDAVP 1-Desamino-8-D-Arginine Vasopressin.

ELISA Enzyme-Linked ImmunoSorbent Assay.

FIM Fisher Information Matrix.

FVIII blood clotting factor VIII.

H High molecular-weight VWF multimers.

HAS Historical Average Subject.

HnonO Healthy subjects not belonging to O ABO group.

HO Healthy subjects belonging to O ABO group.

HY Healthy subjects (by merging subjects from HO and HnonO classes).

KNN *K*-Nearest Neighbors.

L Low molecular-weight VWF multimers.

LDA Linear Discriminant Analysis.

LOF lack-of-fit test.

SSWR Sum of Squared Weighted Residuals.

SUL Super Ultra Large molecular-weight VWF multimers.

UL Ultra Large molecular-weight VWF multimers.

VWD von Willebrand Disease.

VWF von Willebrand Factor.

VWF:Ag von Willebrand Factor antigen.

VWF:CB von Willebrand Factor collagen binding activity.

Roman Symbols

BW body weight (kg).

$C(l|g)$ misclassification cost, the cost of classifying an observation as l when its true class is g .

CL clearance (ml/kg/h).

d_{\cos} cosine distance.

ER error rate for a classification.

$f_g(\mathbf{x})$ Gaussian density function for class g .

\hat{l} predicted class label after classification.

N_G number of classes.

n_{gk} element of the confusion matrix, represent object belonging to class g assigned in class k .

N_p number of predictors of the classification.

N_{sp} number of sampling points.

N_{θ} number of parameters.

N_y number of measured responses.

$\hat{P}(g|\mathbf{u})$ posterior probability for subject \mathbf{u} to belong to class g for.

Pr_g Precision of the g -th class to not include subjects from other classes.

Q total amount of VWF released (IU/kg).

R ratio between VWF:CB and VWF:Ag at the basal state.

SE standard error.

s_{ij} i, j -th element of the inverse matrix of the experimental errors.

Sn_g sensitivity of the g -th class to recognize correctly its subjects.

Sp_g specificity of the g -th class to reject subjects from other classes.

t_{\max} lag time (min).

V_d volume of distribution (dL).

W weight of the distance in the KNN method.

x^L number of low molecular-weight multimer units of VWF.

x^{UL+H} number of ultra large and high molecular-weight multimer units of VWF.

y experimental data of the response.

\bar{y} historical data of the response.

\hat{y} predicted response obtained from the model.

Greek Letters

γ heteroscedastic factor.

ω parameter of heteroscedastic variance.

σ_{err} constant standard deviation for the generation of in-silico experiments.

σ_e heteroscedastic variance of the pool.

π_g prior probability of class g .

τ test duration (min).

\mathcal{Y} factor for the calculation of the posterior probability in the KNN, equal to one if the subject belongs to that class and zero otherwise.

χ^2 chi-square value, used in lack-of-fit test.

Matrixes and Arrays

\mathbf{H}_{θ}^* dynamic information matrix $[N_{\theta} \times N_{\theta}]$.

\mathbf{H}_{θ} discrete dynamic information matrix $[N_{\theta} \times N_{\theta}]$.

$\mathbf{Q}_{i|k}$ dynamic sensitivity matrix for the i -th measured response in the k -th experiment $[N_{\text{sp}} \times N_{\theta}]$.

Σ covariance matrix of the class $[N_p \times N_p]$.

$\boldsymbol{\varphi}$ design vector $[-]$.

$\boldsymbol{\theta}$ array of parameters of the model $[1 \times N_{\theta}]$.

REFERENCES

- [1] F. Rodeghiero, G. Castaman, E. Dini, *Blood*, **1987**, 69, 454–459.
- [2] E. J. Werner, E. H. Broxson, E. L. Tucker, D. S. Giroux, J. Shults, T. C. Abshire, *J. Pediatr.*, **1993**, 123, 893–898.
- [3] V. Kumar, A. K. Abbas, J. C. Aster, *Robbins and Cotran; pathologic basis of disease*, 9 ed., Elsevier Health Science, Philadelphia, U.S.A., **2015**.
- [4] J. E. Sadler, P. M. Mannucci, E. Berntop, N. Bochkov, V. Boulyjenkov, D. Ginsburg, D. Meyer, I. Peake, F. Rodeghiero, A. Srivastava, *Thromb. Haemostasis*, **2000**, 84, 160–174.
- [5] D. Lillicrap, *Thromb. Res.*, **2007**, 120, S11–S16.
- [6] L. Gallinaro, M. G. Cattini, M. Szutukowska, R. Padrini, F. Sartorello, E. Pontara, A. Bertomoro, V. Daidone, A. Pagnan, A. Casonato, *Blood*, **2008**, 111, 3540–3545.
- [7] A. Casonato, E. Pontara, F. Sartorello, M. Cattini, M. Sartori, R. Padrini, A. Girolami, *Blood*, **2002**, 99, 180–184.
- [8] N.H.L.B.I., *The Diagnosis, Evaluation and Management of von Willebrand Disease*, NIH Publication No. 08-5832, US Department of Health and Human Services - National Institutes of Health - National Heart, Lung and Blood Institution, Bethesda, USA, **2007**. URL: <https://www.nhlbi.nih.gov/files/docs/guidelines/vwd.pdf>.
- [9] F. Galvanin, M. Barolo, R. Padrini, A. Casonato, F. Bezzo, *AIChE J.*, **2014**, 60, 1718–1727.
- [10] A. Casonato, L. Gallinaro, M. G. Cattini, E. Pontara, R. Padrini, A. Bertomoro, V. Daidone, A. Pagnan, *Haematologica*, **2010**, 95, 1366–1372.
- [11] Z. M. Ruggeri, P. M. Mannucci, R. Lombardi, A. B. Federici, T. S. Zimmerman, *Blood*, **1982**, 59, 1272–1278.
- [12] J. E. Sadler, *Thromb. Haemostasis*, **2005**, 3, 1702–1709.
- [13] M. Ferrari, F. Galvanin, M. Barolo, V. Daidone, R. Padrini, F. Bezzo, A. Casonato, “A mechanistic model to quantify von Willebrand factor release, survival and proteolysis in Type 2 von Willebrand disease”, *J. Thromb. Haemost.*, **submitted**.
- [14] P. J. Lenting, E. Westein, V. Terraube, A. S. Ribba, E. G. Huizinga, D. Meyer, P. G. de Groot, C. V. Denis, *J. Biol. Chem.*, **2004**, 279, 12102–12109.

- [15] D. Menache, D. Aronson, F. Darr, R. Montgomery, J. Gill, C. Kessler, J. Lusher, P. Phatak, A. Shapiro, A. Thompson, G. White, *Brit. J. Haematol.*, **1996**, 94, 740–745.
- [16] F. Galvanin, C. C. Ballan, M. Barolo, F. Bezzo, *J. Pharmacokinet Phar.*, **2013**, 40, 451–467.
- [17] F. Galvanin, A. Monte, A. Casonato, R. Padrini, M. Barolo, F. Bezzo, “Towards model-based diagnosis of von Willebrand disease”, *24th European Symposium on Computer Aided Process Engineering - ESCAPE 24*, J.J. Klemeš, P.S. Varbanov, P.Y. Liew, Eds. Budapest, Hungary, 15–18 June **2014**, 583–588.
- [18] L. C. Zullo, *Computer aided design of experiments. An engineer approach*, PhD Thesis, University of London, London, UK, **1991**.
- [19] F. Pukelsheim, *Optimal design of experiments*, John Wiley & Sons, New York, U.S.A, **1993**.
- [20] M. Sarkar, T. Y. Leong, “Application of K-nearest neighbors algorithm on breast cancer diagnosis problem”, In: *Proceedings of the AMIA Symposium*, American Medical Informatics Association, **2000**, 759–763.
- [21] T. M. Cover, P. E. Hart, *IEEE T. Inform. Theory*, **1967**, 13, 21–27.
- [22] A. R. Webb, *Statistical pattern recognition*, 2 ed., John Wiley & Sons, Malvern, UK, **2002**.
- [23] D. Ballabio, R. Todeschini, “Multivariate classification for qualitative analysis”, *Infrared spectroscopy for food quality analysis and control*, D. W. Sun Eds., Elsevier: Burlington, MA, USA, **2009**, 83–104.
- [24] D. G. Morrison, *J. Marketing Res.*, **1969**, 6, 156–163.
- [25] D. J. Burn, G. V. Sawle, D. J. Brooks, *J. Neurol. Neurosur. Ps.*, **1994**, 57, 278–284.
- [26] C. H. Miller, J. B. Graham, L. R. Goldin, R. C. Elston, *Blood*, **1979**, 54, 137–145.
- [27] R. A. Fisher, *Ann. Hum. Genet.*, **1936**, 7, 179–188.
- [28] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning; Data Mining, Inference and Prediction*, 2 ed., Springer, New York, USA, **2008**.
- [29] A. Casonato, A. Steffan, E. Pontara, A. Zucchetto, C. Rossi, L. De Marco, A. Girolami, *Thromb. Haemostasis*, **1999**, 81, 224–228.

Figure Captions

1	Historical average subject data for each class for responses of (a) VWF:Ag and (b) VWF:CB. Error bars represent SE of the mean.	4
2	Structure of the new pharmacokinetic model, showing the compartments involved and the available measurements.	5
3	Three-step procedure for in-silico experiment assessment.	8
4	Profiles of the trace of the information matrix over time, computed for each historical average subject.	9
5	Total amount of information for different sampling approaches, different number of samples and different test duration. The results refer to the HnonO historical average subject.	10
6	Pairwise representation of the three indices Q , CL and R , characterizing a given subject, for all subjects considered in this study. Subjects are grouped into three classes thanks to medical examination: healthy (HY), 2B and Vic.	12
7	LDA classification. Representation of boundaries between each couple of classes obtained with all subjects used in model calibration. In this figure is clear the effect of misclassification costs $C(HY 2B)$ and $C(HY Vic)$: increasing them (the direction of the arrows), the bounds are shifted towards healthy region. Since $C(2B Vic)$ was not modified, this boundary is not affected.	20
8	LDA classification. Representation of uncertainty boundaries built with leave-one-out procedure with a $C(HY 2B) = C(HY Vic) = 10$	20

Table Captions

1	Results of parameter estimation task for each historical average subject of the class. The cases in which lack of fit test failed are highlighted in bold.	7
2	Values of σ_{err} under the assumption of constant variance model for the generation of errors, for each class of subjects. These values were calculated and rounded to the closest integer, from the standard deviation of the data of the historical average subject with respect to the predicted responses profiles.	8
3	Heteroscedastic parameters for the four classes considered in classification Method 1. The parameters were obtained from regression on real data.	13
4	Results of model-based diagnosis in terms of SSWR on four unknown subjects. Between parenthesis the real membership class, in boldface the lowest value suggesting the diagnosis.	18

5	KNN classification. Results in terms of performance indices in the cases of no misclassification cost ($C(\text{HY} 2\text{B}) = C(\text{HY} \text{Vic}) = 1$) and with a value of 5.	19
6	LDA classification. Results in terms of performance indices in the cases of no misclassification cost ($C(\text{HY} 2\text{B}) = C(\text{HY} \text{Vic}) = 1$) and with a value of 10.	21